Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia IUTAM 5 (2012) 160 – 164
Procedia IUTAM
www.elsevier.com/locate/procedia

IUTAM Symposium on 50 Years of Chaos: Applied and Theoretical

Network Analysis Based on Statistical-Thermodynamical Formalism

Syuji Miyazaki[a], Taro Takaguchi[b]

[a]Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
[b]Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan

Abstract

Random walk on a graph is analyzed on the basis of the statistical-thermodynamics formalism to find phase transitions in network structure. Each phase can be related to a characteristic local structure of the network. For this purpose, the generalized transition matrix or the generalized Frobenius-Perron operator is introduced, whose largest eigenvalue yields statistical structure functions. The weighted visiting frequency related to the Gibbs probability measure, which turn out to be useful to extract characteristic local structures, is obtained from the inner product of the right and left eigenvectors corresponding to the largest eigenvalue. An algorithm to extract the characteristic local structure of each phase is also suggested based on this weighted visiting frequency.

Keywords: large deviation; rate function

1. Introduction

One of the most remarkable points about deterministic chaos is the duality consisting of irregular dynamics and fractal structure of the attractor in the phase space. Amplifying this relationship between dynamics and geometry, we will try to construct dynamics corresponding to a network structure. A network or graph can be represented by a transition matrix. On the other hand, temporal evolution of a chaotic piecewise-linear one-dimensional map with Markov partition can be governed by a Frobenius-Perron matrix. Both transition matrices and Frobenius-Perron matrices belong to a class of stochastic matrices sharing the same mathematical properties that the maximum eigenvalues is equal to unity and that all the elements of the corresponding right and left eigenvectors are always real. These eigenvectors yield the probability density of visiting a subinterval of the map or a site of the network, which is commercially valuable information in the field of the WWW [1]. Relating these two matrices to each

 doi:10.1016/j.piutam.2012.06.020

other, we are able to represent the structure of the network as a dynamical system. Once we relate the network to chaotic dynamics, several approaches to deterministic chaos can be also applied to graph or network.

In chaotic dynamical systems, local expansion rates which evaluate an orbital instability fluctuate largely in time, reflecting a complex structure in the phase space. Its average is called the Lyapunov exponent, whose positive sign is a practical criterion of chaos. There exist numerous investigations based on large deviation statistics in which one considers distributions of coarse-grained expansion rates (finite-time Lyapunov exponent) in order to extract large deviations caused by non-hyperbolicities or long correlations in the vicinity of bifurcation points [2]. In general, statistical structure functions consisting of weighted averages, variances, and these partition functions as well as fluctuation spectra of coarse-grained dynamic variables can be obtained by processing the time series numerically. In the case of the piecewise-linear map with Markov partition, we can obtain these structure functions analytically. This is one of the reasons why we correspond a network to a piecewise-linear map. We herein try to apply an approach based on an weighted visiting frequency corresponding to the Gibbs probability measure and large deviation statistics in the research field of chaotic dynamical systems to network analyses [3].

2. Rate-function based network analysis

Let us consider the following four graphs. Each node of the graphs has a time-independent internal degree-of-freedom, called node-dependent quantity hereafter, which can be regarded as a user behavior interested or uninterested in a commercial product in the context of marketing research on a social networking service (SNS) site. In other instances, we think of a website inclusive or exclusive of a specific search term, which can be distinguished with the node-dependent quantity 0 or 1. In the following figure, the node-dependent quantity is shown as blue or red circles. In this example, it is discrete, but a continuous node-dependent quantity can also be considered.

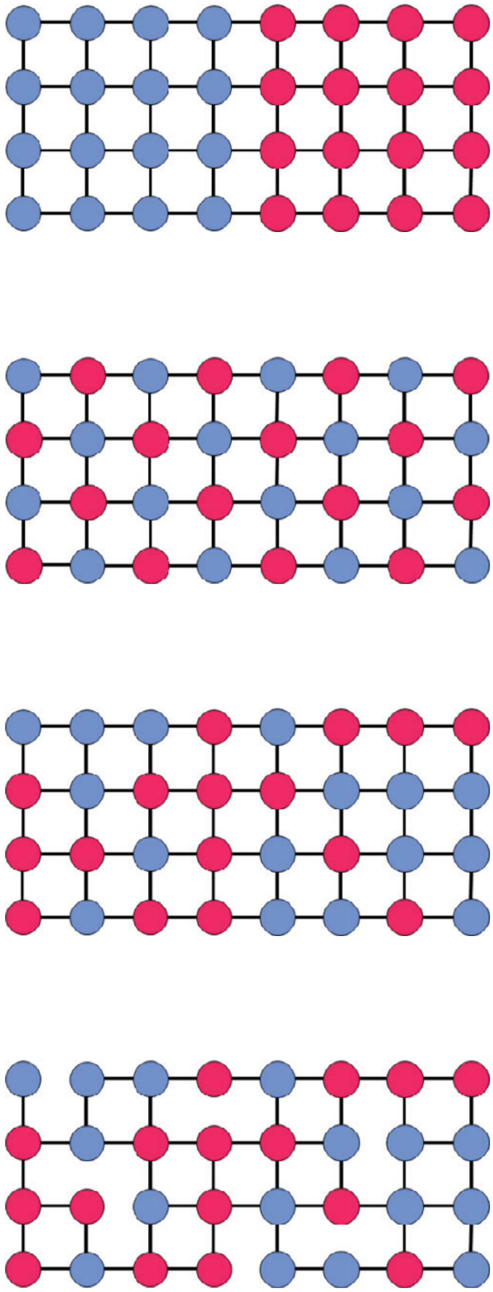

Fig. 1. Four example graphs (a)-(d) from top to bottom.

First three graphs are square lattices. Some links are removed from the square lattice for the fourth graph. The arrangements of blue and red circles are identical for the third and the fourth graph. The link structure and the arrangement of the node-dependent quantity can be visually distinguished for these four graphs. The larger becomes network size, the more difficult becomes such a distinction. We hope that random walk on a graph is a promising analytical approach. We arrange the node-dependent quantity of the node which is visited by the random walker in temporal order. The large deviation statistics of this random time series can characterize the link structure and the distribution of the node-dependent quantity simultaneously.

For stationary discrete-time signals u'_j (j=1,2,…), we consider the following local average over n steps $u_n=(u'_1+u'_2+\ldots+u'_n)/n$. As n goes to infinity, u_n coincides with the long-time average $\langle u \rangle$. For a large but finite n, u_n fluctuates and distributes. Let the distribution function be $P(u,n)$. Even for random or chaotic time series, there exists a characteristic time scale n_c of correlation decay. $P(u,n)$ is proportional to $\exp(-nS(u))$ for $n >> n_c$. where $S(u)$ is called fluctuation spectrum or rate function [4,5].

For the first graph, random time series consisting of 0 and 1 is observed between two extreme cases: …00…00… and …11…11… The rate function of the node-dependent quantity $S(u)$ is continuous and convex downward, which has the minimum value 0 at the long-time average 1/2: $S(1/2)=0$. For the second graph, all the trajectories of the random walker yield periodic time series 010101… The rate function is given by the single point (1/2,0). Though the arrangements of the node-dependent quantity are identical, the rate functions are different for the second and the fourth graph reflecting the different link structures.

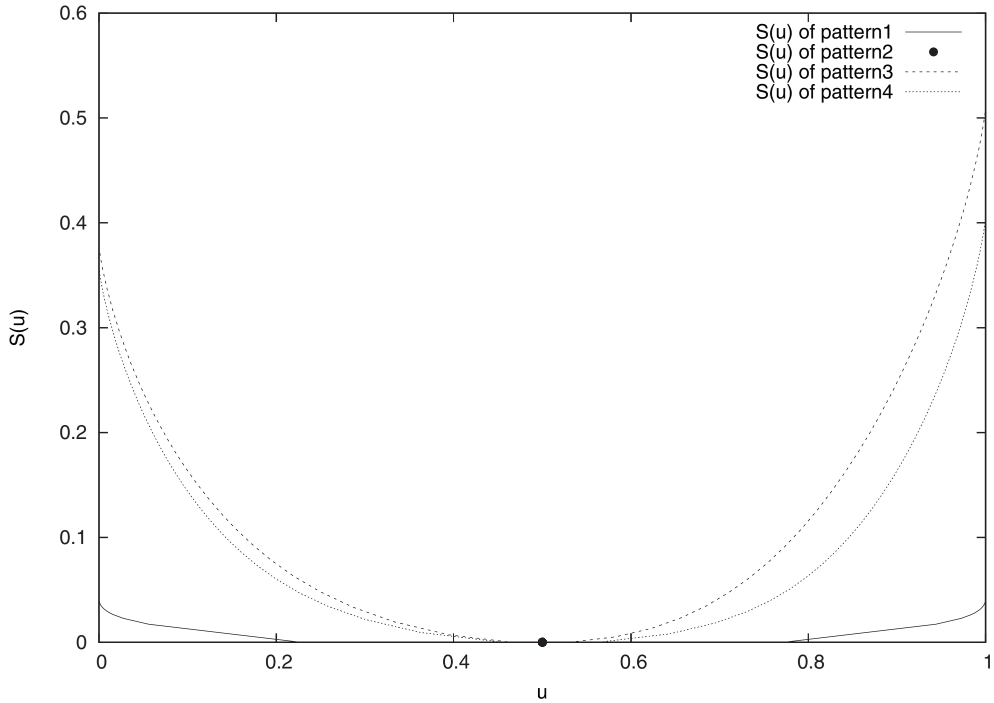

Fig. 2. Four rate functions of the node-dependent quantity corresponding to Fig.1. (a) solid line (b) point (1/2,0) (c) dashed line (d) dotted line.

3. Applications

We applied our analysis based on the statistical-thermodynamics formalism to a real social networking service (SNS). Our analysis object is in such a way constructed that we choose all users within second-neighbor distance from a specific user belonging to the largest SNS in Japan called *mixi* [6]. Let us regard user as node, *my-mixi* relation indicating a friendship on the SNS as undirected link, so that we have an undirected graph with 2271 nodes, among which 11559 undirected links exist. The *mixi* users specify some *keywords* such as *fashion*, *cooking* as their matters of concern. For a fixed *keyword*, we assign the node-dependent quantity $u'=1$, when the node (user) chooses the *keyword* and $u'=0$ otherwise. Random walk on the object graph yields random sequence of 0 and 1 of u'. We analyzed this random time series on the basis of the statistical-thermodynamics formalism to find phase transitions in network structure. Each phase can be related to a characteristic local structure of the network [7].

4. Concluding Remarks

The study introduced in the previous section is performed for a realistic but a specific network. The relationship between the phase of the phase transition and the local structure of the network must be further analyzed. In that purpose, Tanaka studies the Watts-Strogatz model [8] based on the statistical-thermodynamical formalism [9].

Acknowledgements

We thank Kenta Tanaka for his assistance in our numerical analysis. This work was supported by a Grant-in-Aid for Scientific Research (c) (No. 20540376). This presentation is dedicated to Professor Emeritus Hazime Mori, Kyushu University, who passed away suddenly on December 28, 2011.

References

[1] Brin S, Page L. *Comput Netw* 1998;**30**:107-117.
[2] Mori H, Kuramoto Y. *Dissipative Structures and Chaos*. Berlin: Springer; 1998.
[3] Miyazaki S.*Forma* 2007; **22**:141 and references cited therein.
[4] Fujisaka H. *Prog Theor Phys* 1983;**70**:1264.
[5] Fujisaka H. *Prog Theor Phys* 1984;**71**:513.
[6] http://mixi.jp/
[7] Takaguchi T, Ejima K, Miyazaki S. *Prog Theor Phys* 2010;**124**:27.
[8] Watts DJ, Strogatz SH. *Nature* 1998;**393**:440.
[9] Tanaka K. *Thesis (written in Japanese, Faculty of Engineering, Kyoto Univ.)* 2012.